

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/09, 15/12, 15/33, 15/64	A1	(11) International Publication Number: WO 96/09378 (43) International Publication Date: 28 March 1996 (28.03.96)
(21) International Application Number: PCT/US95/11511 (22) International Filing Date: 8 September 1995 (08.09.95) (30) Priority Data: 08/324,243 19 September 1994 (19.09.94) US (71) Applicant: THE GENERAL HOSPITAL CORPORATION [US/US]; 55 Fruit Street, Boston, MA 02114 (US). (72) Inventor: SEED, Brian ; Apartment 5J, Nine Hawthorne Place, Boston, MA 02114 (US). (74) Agent: LECH, Karen, F. ; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02114 (US).		(81) Designated States: AU, BG, BR, BY, CA, CN, CZ, FI, HU, JP, KR, MX, NO, NZ, PL, RO, RU, SG, SI, UA, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: OVEREXPRESSION OF MAMMALIAN AND VIRAL PROTEINS		
(57) Abstract The invention features a synthetic gene encoding a protein normally expressed in mammalian cells wherein at least one non-preferred or less preferred codon in the natural gene encoding the mammalian protein has been replaced by a preferred codon encoding the same amino acid.		

- 55 -

1. A synthetic gene encoding a protein normally expressed in mammalian cells wherein at least one non-preferred or less preferred codon in the natural gene encoding said mammalian protein has been replaced by a preferred codon encoding the same amino acid.

2. The synthetic gene of claim 1 wherein said synthetic gene is capable of expressing said mammalian protein at a level which is at least 110% of that expressed by said natural gene in an in vitro mammalian cell culture system under identical conditions.

3. The synthetic gene of claim 1 wherein said synthetic gene is capable of expressing said mammalian protein at a level which is at least 150% of that expressed by said natural gene in an in vitro cell culture system under identical conditions.

4. The synthetic gene of claim 1 wherein said synthetic gene is capable of expressing said mammalian protein at a level which is at least 200% of that expressed by said natural gene in an in vitro cell culture system under identical conditions.

5. The synthetic gene of claim 1 wherein said synthetic gene is capable of expressing said mammalian protein at a level which is at least 500% of that expressed by said natural gene in an in vitro cell culture system under identical conditions.

6. The synthetic gene of claim 1 wherein said synthetic gene is capable of expressing said mammalian protein at a level which is at least ten times that expressed by said natural gene in an in vitro cell culture system under identical conditions.

- 56 -

7. The synthetic gene of claim 1 wherein at least 10% of the codons in said natural gene are non-preferred codons.

8. The synthetic gene of claim 1 wherein at least
5 50% of the codons in said natural gene are non-preferred codons.

9. The synthetic gene of claim 1 wherein at least
50% of the non-preferred codons and less preferred codons
present in said natural gene have been replaced by
10 preferred codons.

10. The synthetic gene of claim 1 wherein at
least 90% of the non-preferred codons and less preferred
codons present in said natural gene have been replaced by
preferred codons.

11. The synthetic gene of claim 1 wherein said
15 protein is a retroviral or lentiviral protein.

12. The synthetic gene of claim 11 wherein said
protein is an HIV protein.

13. The synthetic gene of claim 12 wherein said
20 protein is selected from the group consisting of gag,
pol, and env.

14. The synthetic gene of claim 13 wherein said
protein is gp120 or gp160.

15. The synthetic gene of claim 1 wherein said
25 protein is a human protein.

- 57 -

16. A method for preparing a synthetic gene encoding a protein normally expressed by mammalian cells, comprising identifying non-preferred and less-preferred codons in the natural gene encoding said protein and
5 replacing one or more of said non-preferred and less-preferred codons with a preferred codon encoding the same amino acid as the replaced codon.

Synqpl20mn

1 CTGAGATCC ATTGTGCTCT AAAGGAGATA CCGGCGCAGA CACCCCTACC
51 TCGGCTGCTC AGCTGCGCAG GCTGAGGCAA GAGAAGGCCA GAAACCATGC
101 CCAATGGGCTC TTTCGAACCG CTGGCCACCT TGTACCTGCT GGGGATGCTG
151 GTGGCTTCCG TTCTAGCCAC CGAGAAGCTG TGGGTGACCG TGTACTACCG
201 CGTGCCCTGT TGAAGGAGG CCACCAACCAC CTTGTTCTGC GCGAGCGAGC
251 CCAAGGCTTA CCAACCCAG GTGCACAACG TGTGGGCCAC CCAAGGCTGC
301 GTGGCCACCG ACCCCAAACC CAGGAGGCTG GAGCTCTGTA AGGTGACCGA
351 GAATTCACAC ATGTGGAAGA ACAACATGCT GGAGCAGATG CATGAGGACA
401 TCATCAGCTT GTGGGACCAAG AGCCTGAAGC CTTGCTGTGA GCTGACCCCG
451 CTGTGCTGTA CTTGAACCTG CACCTACCTG AGGAACACCA CCAACACCA
501 CAACAGCACC GCAACAAACA ACAGCAACAG CGAGGGCACC ATCAAGGGCG
551 CGGAGATGAA CACTGCAGC TTCAACATCA CCAACAGCAT CCGCGACAAG
601 ATGCAGAAGG ATTACGCTCT GCTGTACAAG CTGGATATCG TGAGCATCGA
651 CAACGACAGC ACCAGCTACC GCTGTATCTC CTGCAACACC AGCTGTATCA
701 CCAAGGCTGT GCGCAAGATC AGCTTCGAGC CCACTCCCAT CCACTACTGC
751 GCGCGCGCGG GCTTGGCTAT CTTGAAGTGC AACGACAAGA AGTTCAGCGG
801 CAAGGGCAGC TCAAGAAAGC TGAGCAGCTT GCAGTGCACC CACGSCATCC
851 GCGCGCTGCT GAGCACCCAG CTCTTCTGTA ACGGCAGCTT GCGCGAGGAG
901 GAGGTGCTGA TCCGAGCGCA GAATTCACCC GACAACGCCA AGACCATCAT
951 CTTGCACCTG AATGAGAGCG TGCAGATCAA CTGCACGCTT CCAACTACA
1001 ACAAGGSCAA GCGCATCCAC ATCGGCGCGG GCGCGCTT CTACACCACC
1051 AAGAACATCA TCGGCACCAT CCGCCAGGCG CACTGCAACA TCTCTAGAGC
1101 CAAGTGGAAAC GACACCTGTC GCGAGATGCT GAGCAAGCTG AAGGAGCAGT
1151 TCAAGAACAAC GACCATCTGT TTCAACCAGA GCAGCGCGCG GACCCCGAG
1201 ATCTGTATGC ACAGCTTCAA CTGCGCGCGG GAATTTCTCT ACTGCAACAC
1251 CAGCGCGCTG TTCAACAGCA CTTGGAACCG CAACAAACAC TGAACAACA
1301 CCACCGGCGAG CAACAACAAT ATTACCTTCC AGTGAAGAT CAAGCAGATC
1351 ATCAACATGT GCGAGGAGGT GCGCAAGGCG ATGTACGCGC CCGCATCGA
1401 GCGCCAGATC CGGTGCAGCA GCAACATCAC CCGTCTGCTG CTGACCCCGG
1451 ACGCGCGCAA GGCACCGAC ACCAAGGACA CCGAAATCTT CCGCGCGCG

FIG 1
(SHEET 1 OF 4)

1501 GCGGCGSACA TCGCGGACAA CTGAGAGTCT GAGCTGTACA AGTACAAGGT
1551 GGTGACGATC GAGCGGCTGG GCGTGGGCCC CAGCAAGGCC AAGCGCGGCG
1601 TGGTGCAGCG CAGGAAGGCG TAAAGCGGCG GC (SEQ ID NO:34)

3/12

Synqp160mn

1 AUGGAGAAGC TTGCGGTGAC CTTGTACTAC GCGTGTCTGCG TTGCGAAGGA
51 GCGGACGAGC AGCGTCTTCT GCGGACGCGA GCGGAAGGCG TACGACACCG
101 AUGTGCAGAA CTTGTGCGCG AGCGAGGCTT GCGTGTGCGC GCGCGCGAAC
151 GCGGACGAGC TTGAGCTTCT GAGCTGTGAC GAGAACTTCA ACATGTGGAA
201 GAGCAACATC CTGAGGAGCA TCGATGAGGA CATCATCAGC CTGTGGGAGC
251 AGAGCGTGAA GCGGTCTCTG AGCTGTGCGC CCGTGTGCGT GAGCGTCAAC
301 TCGACGAGT TTAGGAGAC GAGCAACAGC AAGAACAGCA CCGCGAGCA
351 GAGCAGCAAC AGCGAGGCGA CATGTAGGCG GCGGACGATC AAGAACTGCA
401 GCTTCAACAT CAGCAGGAGC ATCGCGTACA AGATCGAGCA GAGTACGCGC
451 TGTGTGTACA AGCTGATAT CTTGAGGATC GAGCAAGGCA GAGCGAGCTA
501 GCGGTGTCTG TGTGTGACA GAGCGGTCTT GAGTACGCGC TGTGTGAGCA
551 TCGCTTCTCA GCGGTCTCTG AGCTGTGACT GCGGTCTCTG GCGGTCTCTG
601 ATCTGTGACT GAGCGAGCA GAGTGTGAGC GCGAGGCGCA CTGTCAAGAA
651 CTGTGAGCAG TGTGAGTCA GAGCGGCGAT GCGGTCTCTG CTGTGAGCAG
701 AGCTGTCTCT GAGTGTGAGC CTGTGTGAGC AGAGGCTCTT CATGTGAGC
751 GAGTGTGAGC GAGCGAGCAG GAGGAGGATC ATCTGTGAGC TGAATGAGAG
801 GAGGAGGATC AGCTGTGAGC CTGTGTGAGC GAGGAGGAGC AGGCGAGTCT
851 AGATGTGTCT GCGGTCTCTG TGTGTGAGCA GAGGAGGATC CATGTGAGC
901 ATGTGTGAGC GCGGTCTCTG CATGTGTAGA GCGAGGAGCA AGGAGAGCTT
951 GCGGAGGATC CTGTGAGGAT TGAAGGAGCA GGTCAAGAAC AGAGGAGTCT
1001 TGTGTGAGCA GAGGAGGAGC GCGGAGGAGC AGATGTGTCT GAGGAGGATC
1051 AGCTGTCTCT GCGGATCTCT GAGTGTGAGC AGGAGGAGC TGTGTGAGC
1101 GAGTGTGAGC GCGGAGGAGC CTGTGTGAGC GAGGAGGAGC AGGAGGAGC
1151 ATATTAGCTT GAGTGTGAGC ATGTGTGAGC GAGTGTGAGC CTGTGTGAGC
1201 CTGTGTGAGC CATGTGTCT GCGGTCTCTG GAGGAGGAGC TGTGTGAGC
1251 GAGGAGGATC AGCTGTCTCT TGTGTGAGC GAGGAGGAGC AGGAGGAGC
1301 AGGAGGAGC GAGGAGGATC TGTGTGAGC GCGGTCTCTG CATGTGTCTCT
1351 AGTGTGAGC CTGTGTCTCT GAGTGTGAGC CTGTGTGAGC TGAAGGAGC
1401 GCGGTCTCTG GAGGAGGAGC TGAAGGAGC CTGTGTGAGC GCGGAGGAGC

FIG. 1
(SHEET 3 OF 4)

1451 GGGCGGCGAT GGGCGGCGTG TTGCTGGGCT TCTTGGGGGC GCGGGGCGAG
1501 ACCATGGGGG GGGCGAGCGT GACCTTGACC GTGAGGGGCG GCTTGGTCTT
1551 GAGCGGCGATC GTGAGGCGAG AGAACAACCT GCTCGGCGCG ATCGAGGCGG
1601 AGCAGCATAT GTTCCAGCTG ACCGTGTGGG GCATCAAGCA GCTCCAGGCG
1651 GCGGTGGCTG CGTGGAGCG GTACCTGAAG GACCAAGCAG TCTTGGGCTT
1701 GTGGGGCTGC TGGGGCAAGC TGATCTGCAC CACGACCGTA CCTTGGAAAG
1751 GCTCGTGGAG CAGCAAGAGC GTGAGCGACA TGTGAAACAA CATGAGCTGG
1801 ATCAGTGGG AGCGCGAGAT CGATAACTAC ACCAGCGTGA TTACAGCTT
1851 GTTGAAGAAAG AGCCAGAGCG AGCAGGAGAA GAACGAGCAG GAGCTGCTGG
1901 AGCTGGACAA CTGGGCGAGC GTGTGAAACT GTTTCGACAT CACCAACTGG
1951 GTGTGTACA TAAAAATCTT CATCATGATT GTGGGGGCGC TGTGGGCTT
2001 GCGATCTGTG TTGCGCTGCG TGAGCATCTT GAACGCGCTG CGCGAGGCTT
2051 ACAGCGCGCT GAGCTGCGAG ACCGCGCGCG GCTTGGCGCG CGGGCGCGAC
2101 GCGCGCGAGG GATCGAGGA GAGGGCGCG GAGCGCGACC GCGACCGAG
2151 GCGAGGCTG GTGAGCGCT TCTGGCGAT CATCTGGCTG GACCTCGCGA
2201 GCTGTCTCT GTTCAAGTAC CACCAAGCGG AGCTGCTCT CATCGCGCGC
2251 GCGATCTGTG AACTGCTAGG CGCGCGCGCG TGGAGGTGC TGAAGTACTG
2301 GTGAACTGCT CTCAGTATT GAGCGAGGA GTTGAAGTCC AGCGCGCTGA
2351 GCTTGTGTA GCGACCGCG ATGCGCGTGG GTGAGGGGAC CGACCGCTG
2401 ATCGAGGTGC TCGAGAGGCG CGGAGGGCG ATCTGCGACA TCGCGAGCG
2451 TATCGCGAG AGGCTCGAGA GCGCGCTCT G (Seq ID NO 35)

5/12

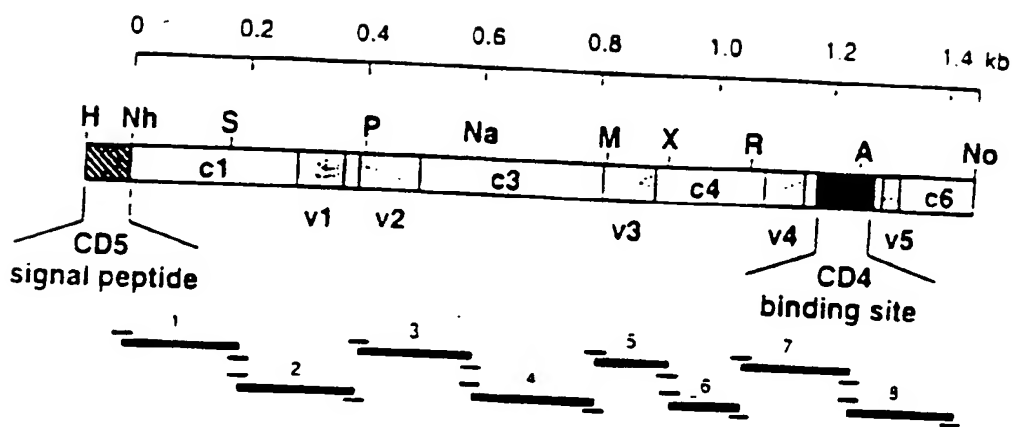


FIGURE 2

6/12

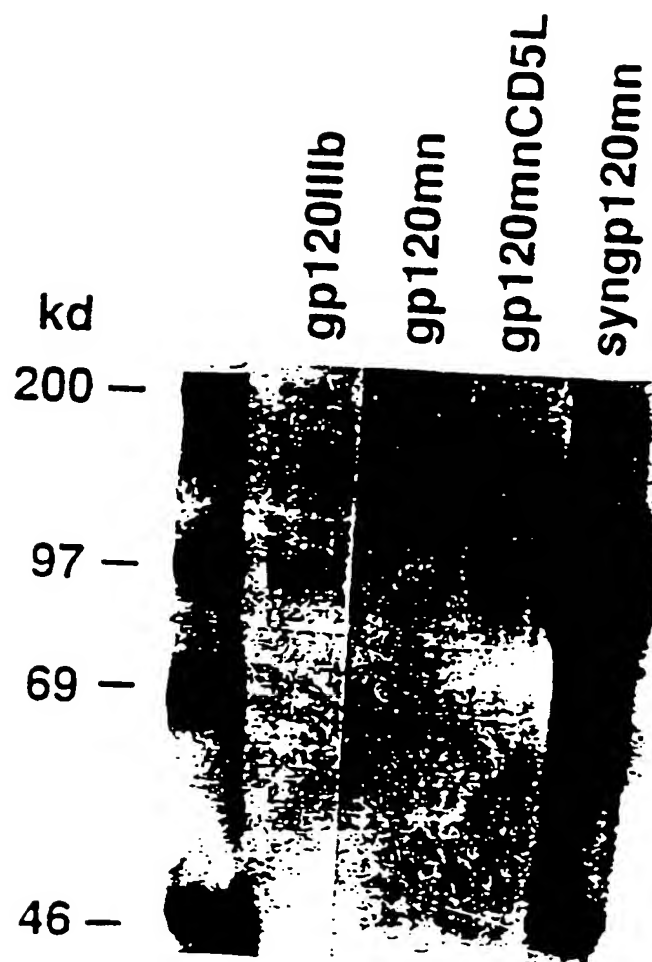


FIGURE 3

7/12

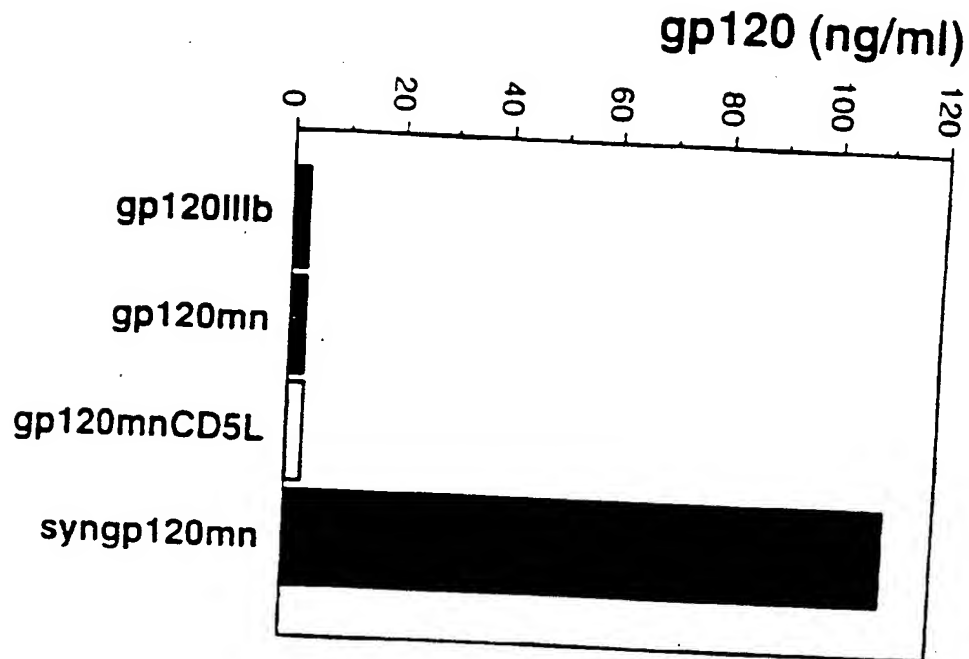


FIGURE 4

8/12

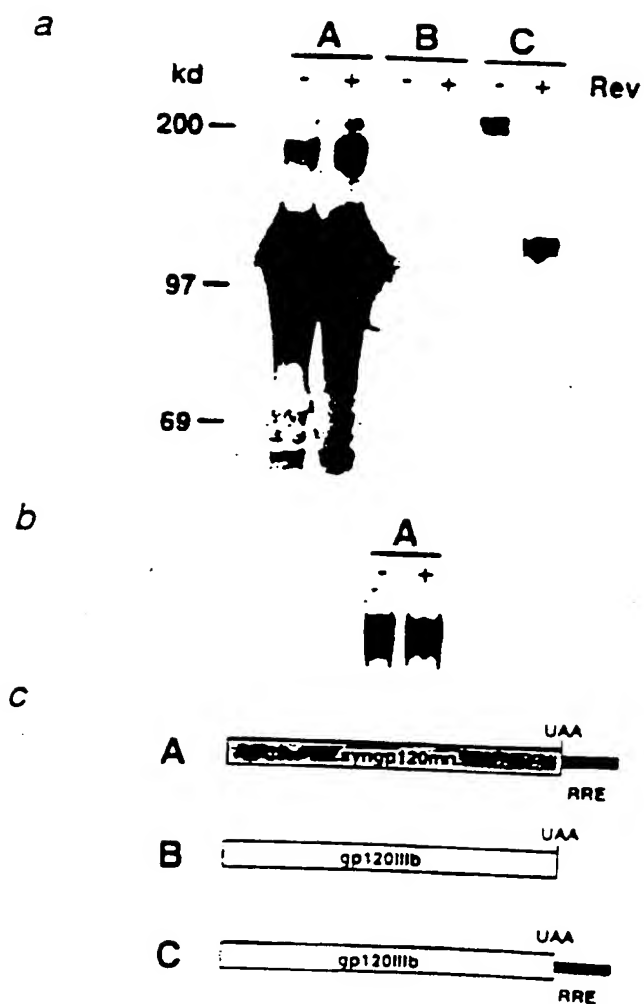


FIGURE 5

FIGURE 6

10/12

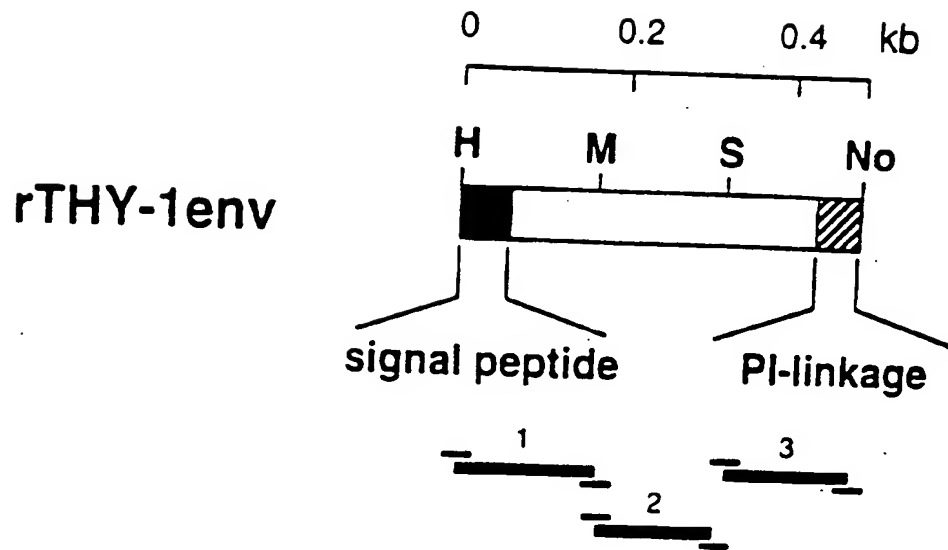


FIGURE 7

11/12

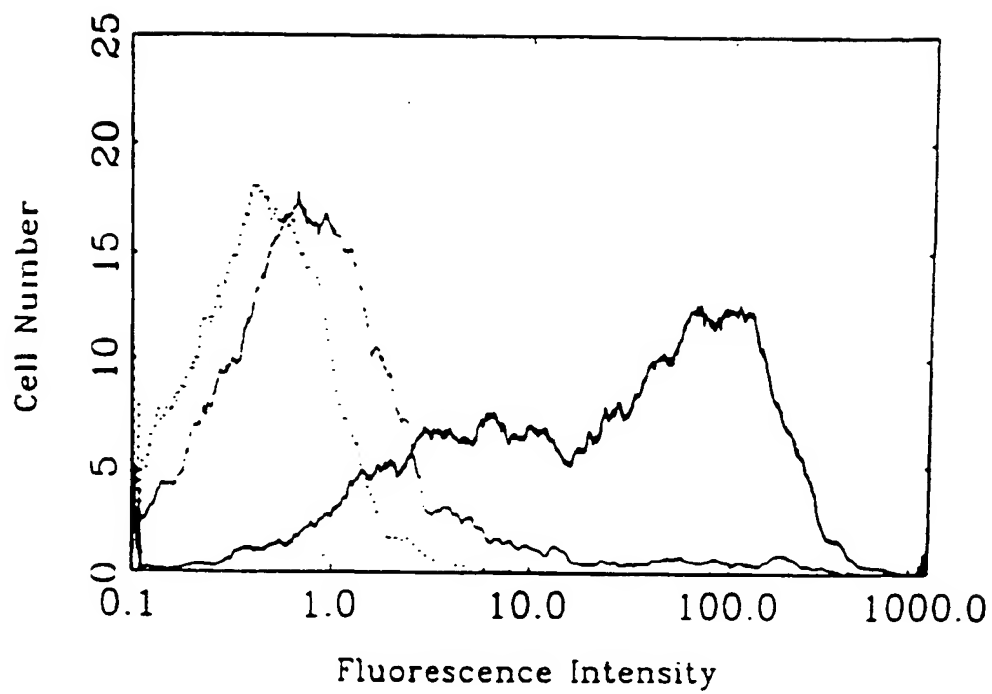


FIGURE 8

12/12

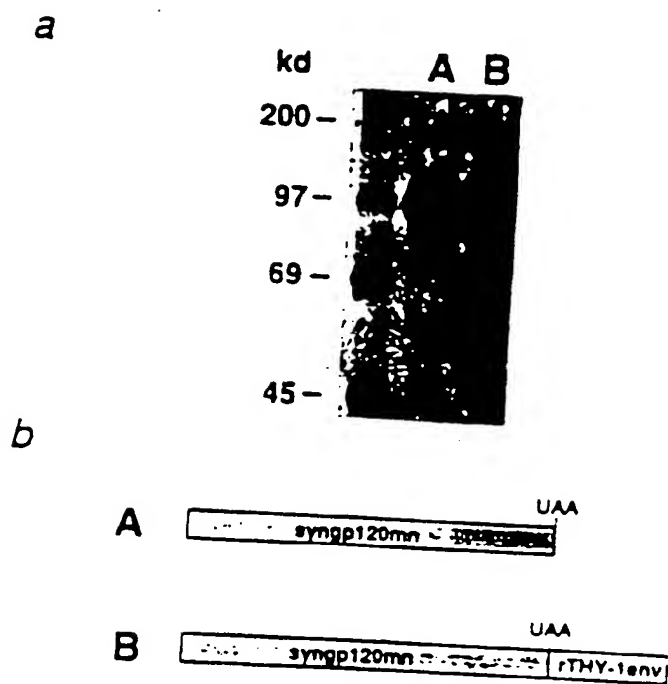


FIGURE 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/11511

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : Please See Extra Sheet.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Please See Extra Sheet.

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,270,171 (CERCEK ET AL.) 14 December 1993, see column 34, lines 32-48.	1-16
Y	Nucleic Acids Research, Volume 18, Number 4, issued 1990, McCarrey, "Molecular evolution of the human Pgk-2 retroposon", pages 949-955, see entire document.	1-16
Y	Japanese Journal of Cancer Research, Volume 80, issued March 1989, Kamiya et al., "Transformation of NIH3T3 Cells with Synthetic c-Ha-ras Genes", pages 200-203, see entire document.	1-16

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	* T	later documents published after the international filing date or priority date and not in conflict with the application but cited to undermine the principle or theory underlying the invention
* A* document defining the general state of the art which is not considered to be of particular relevance	* X	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
* E* earlier documents published on or after the international filing date	* Y	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combinations being obvious to a person skilled in the art
* I* document which may throw doubts on priority claim(s) or which is used to establish the publication date of another citation or other special reasons (as specified)	* A	document member of the same patent family
* O* document referring to an oral disclosure, use, exhibition or other means		
* P* document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

24 OCTOBER 1995

Date of mailing of the international search report

03 NOV 1995

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

JAMES KETTER

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

Int. national application No.
PCT/US95/11511

C (Continuation), DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Nucleic Acids Research, Volume 16, Number 17, issued 1988. Sharp et al., "Codon usage patterns in <i>Escherichia coli</i> , <i>Bacillus subtilis</i> , <i>Saccharomyces cerevisiae</i> , <i>Schizosaccharomyces pombe</i> , <i>Drosophila melanogaster</i> and <i>Homo sapiens</i> ; a review of the considerable within-species diversity", pages 8207-8211, see entire document.	1-16
Y	Proceedings of the National Academy of Sciences USA, Volume 83, issued November 1986. Newgard et al., "Sequence analysis of the cDNA encoding human liver glycogen phosphorylase reveals tissue-specific codon usage", pages 8132-8136, see entire document.	1-16
Y	Gene, Volume 46, issued 1986. Coulombe et al., "Expression of a synthetic human interferon- α gene with modified nucleotide sequence in mammalian cells", pages 89-95, see entire document.	1-16

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/11511

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C12N 15/09, 15/12, 15/33, 15/64

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

536/23.5, 23.72; 435/172.3

B. FIELDS SEARCHED

Minimum documentation searched

Classification System: U.S.

536/23.5, 23.72; 435/172.3

B. FIELDS SEARCHED

Documentation other than minimum documentation that are included in the fields searched:

NONE

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS. MEDLINE EXPRESS